DOCUMENT RESUME

ED 026 130 PS 001 503

By-Stolz, Walter; Bills, Garland

An Investigation of the Standard-Nonstandard Dimension of Central Texan English. Part of the Final Report.

Texas Univ., Austin. Child Development Evaluation and Research Center.
Spons Agency-Office of Economic Opportunity, Washington, D.C.
Report No-OEO-4115
Pub Date Aug 68
Note-51p.
EDRS Price MF-$0.25 HC-$2.65
Descriptors-Adult Characteristics, *American English, Correlation, Distinctive Features, Educational Background, Educational Status Comparison, *Language Research, *Regional Dialects, Social Dialects, Socioeconomic Status, *Sociolinguistics, Speech Habits
Identifiers-Scalogram Analysis

The speech of 23 people in a rural Texas community was studied. The population of the area was over 90 percent white Anglo-Protestant. The subjects varied from illiterate to college educated and ranged in age from 17 to 60 years. Seventeen dialect features were chosen as illustrative of the variations in the use of linguistic forms for people of this geographical area. It was assumed prior to the study that there existed a continuum on which both the speaker and the dialect features could be located. Twelve of the 17 dialect features were analyzed. A correlation was found between dialect score (DS) and amount of education of the subject; the higher the education, the closer the speaker was to the standard English end of the continuum. A similar result was found between DS and socioeconomic level. In another part of the study, one of the dialect features was intensively analyzed for the speech of 12 subjects. A correlation was found between socioeconomic class and the use of "have" or "have got," with the lower classes using "have got." (WD)

ED026130

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PART OF THE FINAL REPORT

to

THE OFFICE OF ECONOMIC OPPORTUNITY

(Contract No. OEO-4115)

CHILD DEVELOPMENT EVALUATION AND RESEARCH CENTER

John Pierce-Jones, Ph.D., Director

The University of Texas at Austin

August, 1968

AN INVESTIGATION OF THE STANDARD-NONSTANDARD DIMENSION OF CENTRAL TEXAN ENGLISH

by

Walter Stolz

and

Garland Bills

PS001503

PART OF THE FINAL REPORT

to

THE OFFICE OF ECONOMIC OPPORTUNITY

(Contract No. OEO-4115)

CHILD DEVELOPMENT EVALUATION AND RESEARCH CENTER

John Pierce-Jones, Ph.D., Director

The University of Texas at Austin

August, 1968

AN INVESTIGATION OF THE STANDARD-NONSTANDARD DIMENSION OF CENTRAL TEXAN ENGLISH

by

Walter Stolz

and

Garland Bills

AN INVESTIGATION OF THE STANDARD-NONSTANDARD DIMENSION OF CENTRAL TEXAN ENGLISH

Walter Stolz

and

Garland Bills

The University of Texas at Austin

ABSTRACT

In Part A linguistic continuum was postulated to run between extreme standard English and nonstandard English. Further, it was assumed that both speakers and dialect indicators could be ordered along it. Twelve dialect indicators for 23 informants were subjected to scalogram analysis. A coefficient of reproducibility of more than .96 was obtained. Correlations between an informant's score on the resulting scale and his socioeconomic status and amount of formal education was found to be highly significant. A suggestion was made to integrate the concepts of the standard-nonstandard continuum with the formal-informal continuum of style. Also, speculations were made about implications of the present study for theoretical linguistics and psycholinguistics.

In Part B, one of the dialect indicators, the use of _have_ versus _have got_, was intensively analyzed in the speech of 12 informants. The frequencies of usage of _have_ and _have got_ were found to

distribute themselves bimodally with lower-class speakers using <u>have got</u> more than 50% of the time and middle-class speakers using the form less than 30% of the time.

PART A

AN INVESTIGATION OF THE STANDARD-NONSTANDARD DIMENSION OF CENTRAL TEXAN ENGLISH

AN INVESTIGATION OF THE STANDARD-NONSTANDARD DIMENSION OF CENTRAL TEXAN ENGLISH

Walter Stolz

Introduction

Much of the work in dialectology and sociolinguistics has been of a purely descriptive nature. At the most primitive level a large amount of labor has gone into the identification of linguistic elements which vary as a function of either geographical location or social groupings. Thus, a procedure such as this divides the populaton of native speakers into a number of classes, such that each class speks a relatively homogeneous brand of English; that is, they all use or do not use the same linguistic elements (which have been identified as dialect indicators). Viewed abstractly, this process is merely that of selecting a set of variables which are then used to define a taxonomy for the useful classification of the members of a population into subgroups (in this case, the population is the speakers of some language say English). Within the behavioral sciences there are many well-known mathematical procedures for doing this sort of task and the purpose of this report will be to use one of them--known as scalogram analysis (Guttman, 1944) to classify the speakers of a given dialect community.

Scalogram analysis is a measurement "model" or "theory" in its own right; that is, it makes certain assumptions about the properties of the data to which it is applied, and if the data do not exhibit these properties, then various descriptive measures of "fit" between the model and the data will indicate that the model must be rejected. Thus, we can think of this project as testing whether or not certain sorts of dialect data have certain logical properties.

Assumptions of the Model

The basic postulation of scalogram analysis is the existence of a single underlying continuum or dimension on which both speakers and dialect features can be located. In this case, we will assume that such a continuum stretches from the extreme version of any nonstandard dialect of English to standard English (SE). Since this study deals with a dialect community in central Texas composed of rural, white, Anglo-Saxon Protestants who speak a nonstandard variety of English, the nonstandard end of the continuum will be labeled CTRWASPNSE. The model assumes that each of the people in this community can be located somewhere along this dimension--probably depending on the relative extent of their contacts with CTRWASPNSE and with SE (or, in particular, the version of SE spoken by educated Central Texans (SCTE). The model also assumes that features which distinguish CTRWASPNSE from SCTE can also be each located at a point on this continuum in the following sense.

We assume that for each of these features a dichotomous variable can be defined, having a value of either 1 or 0 for each speaker as he speaks in a given style. Any given speaker in a given style is assumed to have either the SCTE version of the dialect feature or the CTRWASPNSE version, but not both; if he has the standard variant, he receives a 1 on that variable, and if he has the nonstandard variant, he receives a 0. For example, if a speaker uses double negative constructions, he would be assigned a 0 on a given variable, say double negative, and is further toward the CTRWASPNSE end of the continuum than is any speaker with a 1 on that variable. Therefore, for each variable it is assumed that there exists some point on the continuum such that all people above that point have a 1 on that variable and all those below that point have a 0. This point is often called the <u>cutting point</u> for that variable. Variables are then ordered on the continuum with respect to where their cutting points come in relation to each other, and people are ordered with respect to their patterns of responses on these variables.

If the data conform to all of these assumptions then only a few of the combinatorially possible patterns of 1's and 0's will actually occur in the data for any sample of individuals. If there are k dialect features under consideration, then each speaker must display one of at most $k + 1$ response patterns. As a hypothetical example, suppose $k = 4$ as in Figure 1, with the cutting points for the features

denoted by the integers 1 through 4. These cutting points divide the continuum into five segments, denoted by the roman numerals I through V. The model requires that all speakers occupying the same segment must have the same pattern of 1's and 0's as shown in the example below:

	FEATURE			
Group	1	2	3	4
I	0	0	0	0
II	1	0	0	0
III	1	1	0	0
IV	1	1	1	0
V	1	1	1	1

As the number of features under consideration increases, the assumptions of the model become relatively more stringent. The combinatorially possible patterns increase as k! and the number permitted under the model increases as k + 1; so, as k becomes very large, $\frac{k+1}{k!}$ approaches zero. A more complete description of the logic and methods of scalogram analysis can be found in Torgerson (1957).

To the extent that the distribution of dialect features in a dialect community follow the assumptions given above, several important implications follow. First, it must be recognized that a community of

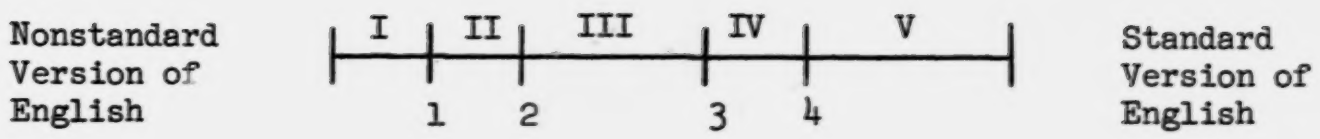

Figure 1. An example of a dialect continuum

PS001503

nonstandard speakers is not a homogeneous group, but contains speakers using all intermediate versions of the language between the extreme standard and extreme nonstandard versions. However, it also implies that the progression from one end of the continuum to the other is an extremely orderly one with respect to the appearance and disappearance of various dialect features. If we know the number of nonstandard features in any given speaker's speech, we can predict exactly which ones they are. Also, if we consider changes in style, for a given speaker, as being shifts in his position along the continuum, it is possible to specify exactly the order in which dialect features will appear and disappear in his speech as he shifts from style to style. This notion is not original here but has been suggested by DeCamp (1968) in the study of Creoles.

Procedure

The speech community which was studied in this investigation was that of an almost totally rural area centered about 20 miles northwest of Austin, Texas. There are no incorporated towns within this region, although the principal school district (Leander ISD) has about 500 children of school age. The only industries are a lime plant and a stone quarry. Some farming is done, but most of the land will only support sheep, goats, and a few cattle. Nearly all of the upper socio-economic class (SEC) members work in Austin and commute daily. The

population of the area is over 90% white Anglo-Protestant and most of those participating in this study were at least second generation Central Texans. The members of the lower SEC have traditionally made their living through cutting the cedar--which covers most of the area--making fence posts, tenant farming, hunting, fishing, and trapping. Most of them are functionally illiterate.

The present study is based on a sample of 23 speakers from this community, 14 males and 9 females. They ranged in SEC from a woman with a master's degree in chemistry (married to a Ph.D. chemist) to an unemployed illiterate laborer who claimed never to have completed the first grade. The informants ranged in age from 17 to 60 years. They included two mother-son pairs, but otherwise were all unrelated to each other. Nineteen of these informants were from a study by Stolz and Legum (1967) and four were first, second, and third grade teachers in the Leander School. Their ages, education, occupation, and SEC are listed in Table 1.

The interview schedule was that used by Stolz and Legum. The speech samples were at least 20 minutes of conversation with one of two interviewers,[1] mainly in free format with the primary objective being to elicit some extended narrative in a relatively informal style from

---

[1]The interviewers were Mr. Stanly Legum and Mr. Robert Berdan, graduate students in linguistics at The University of Texas at Austin. Each conducted roughly half of the interviews using the same schedule.

TABLE 1

LIST OF INFORMANTS

Informant	Sex	Age	Highest Education (in yrs.)	Occupation	ISS
AF	F	60	16	Teacher	32
MJW	F	43	16	Teacher	28
ML	F	39	18	Chemist	26
JP	F	40	17	Teacher	31
RC	F	47	16	Teacher	32
NC	F	28	12	Housewife	57.5
GK	M	34	16	Teacher-Preacher	32
DW	M	34	10	Gas Service Man	58.5
MW	F	38	12	Teacher's Aide	43
EW	M	17	11*	Student	43
WF	M	43	12	RR General Agent	35.5
LB	M	30	12.5	Journeyman Electrician	48
JM	M	28	13	Painting Contractor	39
KF	M	32	16	Teacher	34
BH	M	36	12.5	Army Officer	40.5
GJ	F	29	10	Housewife	58
MVW	F	26	10	Housewife	56
EH	M	39	5	Field Hand, Cedarchopper	80
ENH	M	40	11.5	Dam Operator	48
GF	M	30	9	Truck Driver	65
JC	M	32	5	Truck Driver	71
HM	M	43	1	Cedar Yard Employee	76
JT	M	45	4	Cedarchopper	80

*Still in school

the informant. All interviews were conducted in the informants' homes except for the teachers, who were interviewed on the school premises during school hours.

The Dialect Features

Seventeen dialect features were initially chosen for use in this study. They were not selected in any particularly systematic way, but were identified after many hours of transcribing and studying the speech of the informants. Five of these variables were dropped from the analysis because of low frequency of occurrence, vagueness of definition, etc., however, all seventeen will be identified and discussed below:

1. Double negatives. If an informant used at least one double negative he was given a 0 on this variable, otherwise he was given a 1. Examples were He didn't never go; I usually don't have none; It don't hardly matter.

2. Use of ain't. If an informant used ain't at least once he was given a 0 on this variable, otherwise he was given a 1.

3. Substitution of d for s before -n't. For this variable, only the pronunciation of three words was scored, isn't, wasn't, and doesn't. In CTRWASPNSE isn't becomes idn't, wasn't becomes wadn't, and doesn't becomes doedn't. The mean frequency per interview of these three words together was about five, with only one informant having no

occurrences and one having a single occurrence. If an informant substituted /d/ for /s/ in any occurrence during his interview he was given a zero on this variable, otherwise he was given a 1. The single informant with no occurrences of the words was not scored on the variable.

4. Pronunciation of can't as cain't. This variable was dropped from the analysis for two reasons. First, three of the 23 informants didn't use the word. Second, five used intermediate pronunciations which were extremely difficult to score reliably.

5. Auxiliary and copula deletion. Occasionally the copula is deleted from a sentence (cf. Labov, 1968) as in Most of 'em just things that like to happen but don't. Also, auxiliary verbs are sometimes missing as in People gonna have to start carryin guns. Informants having two or more deletions were scored as 0 on this variable. It should be noted that this threshold was set in a totally post hoc way.

6. Past tense-past participle confusion. In many cases this variable tends to be confounded with 5, as in They done it already. It is impossible to say if have has been omitted, thus making it an auxiliary deletion, or if did has been confused with and replaced by the past participle form, done. Whenever this sort of uncertainty arose, the construction was counted as an instance of 6 rather than of 5. Generally, this variable was scored whenever two or more simple-past verb forms were substituted inappropriately for the past-participle form or vice versa.

7. Have got for have. This variable scored the use of have got for the sense of have meaning "ownership" as in They've got 45 hogs. A separate study of 12 informants (see Appendix A) indicated that the frequency of use of got was essentially bimodal with one mode at about 80-90% and the other at about 5-10%. In a partially post hoc decision the threshold for scoring an informant with a 0 was that more than 30% of the possible opportunities for the use of got actually had it appearing there.

8. Number disagreement between subject and verb. Typical examples of this variable are They was scared plenty and It don't pertain to our work too much. Any instance of this feature in an interview caused a 0 to be assigned to the informant on this variable. Otherwise he received a 1.

9. Use of there is with a plural noun phrase. Apparently separate from variable 8 is the very common use of there is with a plural noun phrase as in There is three of them, there was usually that many, etc. Any use of this form caused the informant to be scored as a 0 on this variable.

10. Substitution of an adjective for an adverb. For example: It was fixed up pretty good for pigs and they found the dime real easy. Two occurrences of such a substitution caused this variable to be scored as a 0.

11. Substitution of that for who as a relative clause introducer. When a restrictive relative clause modifies a human noun phrase,

it is common to use the unmarked clause introducer that instead of the (+ human) introducer who. Thus, the man who went becomes the man that went. Since it appears that everyone used that at least occasionally, any occurrence of who in an informant's speech was counted as a 1. If fewer than five possible chances for this occurred in an interview without a who occurring, the informant was not scored at all. This happened for one informant.

12. Deletion of th- in there or that following a word ending in a stop. Typical CTRWASPNSE pronunciations are like 'at and up 'ere. If either of these deletions occurred twice or more, an informant was given a 0 for the variable; otherwise he was given a 1.

13. Reduction of -ing to -in when used as a bound morpheme. In this case approximately 20 examples were noted for each informant somewhere near the middle of his interview; if two or more were -ing, a 1 was assigned to the informant. Both verb and gerundive forms were counted.[2]

14. Loss of diphthong on words ending in -ire. This variable was abandoned because it appeared not to be dichotomous in nature; that

---

[2]After most of the corpora had been scored, it became apparent that the reduction of -ing and -in was different for gerunds than for progressive verbs. Several speakers regularly pronounced hunting with the -ing when using it in a noun phrase, but reduced the suffix to -in when using it as a verb. This may account for the fact that the -ing variable did not scale as well as the others.

is, all degrees of diphthongization were observed, from barely perceptible to highly pronounced. Thus, the scoring process became highly subjective and probably very unreliable.

15. Substitution of hisself for himself and theirselves for themselves. This variable was not scored because it was found that the third-person reflexive only occurred in about half of the corpora.

16. Substitution of them for those as a determiner. An example would be them dogs for those dogs. This variable was also abandoned because it occurred too infrequently in the corpora.

17. Rising terminal intonation on a declarative utterance. This variable was abandoned for the same reason as were 4 and 14, because all gradations of the feature were observed making dichotomous scoring very difficult.

The Setting of Thresholds

A certain degree of arbitrariness is introduced into this use of scalogram analysis because of the requirement that each informant be given either a 1 or a 0 on each variable. Thus, we can not enter into the analysis each informant's proportion of usage of a given feature; he must be scored categorically on whether he uses the standard version of the feature or not. Since the data usually do not come in this all-or-nothing way, the necessity arose for setting a threshold for each feature--that is, to specify a critical number or proportion of

occurrences such that if any informant exhibited more than that amount of the standard (or for some features, the nonstandard) variant, he was given a 1 (or a 0). Because of our previous inexperience with the dialect, such thresholds were impossible to set *a priori* without a great deal of arbitrariness. Thus, thresholds were set *post hoc* to give the optimal fit between the data and the scalogram model. In all cases except variable #7, the choice was merely between one and two occurrences of a given variant. And this, to a large extent, reflected the relationship of the sizes of the corpora to the baseline frequencies of the various features. For the more infrequent features, the average number of occurrences per interview was small and a threshold of one had to be used; however, two seemed a much more stable threshold.

If an unlimited amount of speech were available from each informant, we would hope that the frequency distribution for a given variant of a given feature would be distinctly bimodal with informants having the standard variant emitting it at one rate (say 70% standard variant to 30% nonstandard) whereas those having the nonstandard variant might have their proportions just reversed. Obviously, then, the threshold for such a feature should be placed at around 50%. Variable #7 was just such a case; with 12 informants studied intensively, six used *got* less than 25% of the time, five used it 70% of the time or more, and one used it 44% of the time. The classification is clear except for the 44% informant and we would hope that with a larger corpus, that informant would join one of the modes.

Results

The results of the scalogram analysis for 12 variables and 23 speakers is given in Table 2. In five cases it was judged that there were too few occurrences of the linguistic feature (either the SCTE or CTRWASPNSE version) to assign a 1 or 0 to the speaker. Two of the five are irrelevant to the analysis in the sense that the analysis makes no prediction about whether they should be 1 or 0. The other three would be predicted to be 0 by the analysis, given more data from the speaker. In most cases below, we will follow through with two possibilities, one that the missing data fit the overall pattern in each case, the other that in each case they do not. The scalogram model predicts that each row of the matrix in Table 2 should have all its 1's (if there are any) to the left of all its 0's. Also, for each column, all the 1's should be above all the 0's. Any deviation from this pattern is considered to be an "error" in the sense that it is a misfit between model and data. Rows and columns that conform to the predicted pattern can be given a score corresponding to the number of 1's in the particular row or column. For rows and columns containing errors, it is not easy to determine what score should be assigned; in the present analysis, the score was assigned which minimized the number of errors overall and the number of CTRWASPNSE forms per speaker. As mentioned above, two scores are given in Table 2, one assuming each missing data cell turned out to

TABLE 2

RESULTS OF THE SCALOGRAM ANALYSES

Informant	Dialect Feature												score	
	ain't	'AuxD	p-pp	DN	ing	s/p	have	th → ø	adj	wadn't	there's	that	max	min
AF	1	1	1	1	1	1	1	1	1	1	1	1	12	12
MJW	1	1	1	1	1	1	1	1	1	1	1	1	12	12
ML	1	1	1	1	1	1	1	1	1	1	1	nd	11	12
JP	1	1	1	1	1	1	1	1	1	1	1	0	11	11
RC	1	1	1	1	1	1	1	1	1	0	1	0	11	11
NC	1	1	1	1	1	1	1	1	1	1	0	1	12	12
GK	1	1	1	1	1	1	1	1	1	1	0	0	10	10
DW	1	1	1	1	1	1	1	1	1	0	0	0	9	9
MW	1	1	1	1	1	1	1	1	1	nd	nd	0	9	11
EW	1	1	1	1	0	1	1	1	1	0	0	0	9	9
WF	1	1	1	1	1	1	1	1	1	0	0	0	9	9
LB	1	1	1	1	0	1	0	1	1	0	0	0	9	9
JM	1	1	1	1	1	1	1	0	0	0	0	0	7	7
KF	1	1	1	1	1	1	1	0	0	0	0	1	7	7
BH	1	1	1	0	1	0	nd*	0	0	0	0	0	5	4
GJ	1	1	1	1	0	0	0	0	0	0	0	0	4	2
MVW	1	1	0	0	0	0	0	0	0	0	0	0	2	1
EH	1	0	0	0	0	0	0	0	0	0	0	0	1	1
ENH	1	0	0	0	0	0	0	0	0	0	0	0	1	0
GF	0	0	0	0	0	0	0	0	nd	0	0	0	0	0
JC	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JT	0	0	0	0	0	0	0	0	0	0	0	0		
Max Score	19	17	16	16	15	14	14	12	12	7	5	2		
Min Score	19	17	16	16	15	14	14	12	12	9	5	3		

*nd = no data

be an error and one assuming that it turned out to fit the row and column pattern already established by the variables around it.

There are several descriptive measures of fit between model and data which have been suggested for evaluating the results of scalogram analysis (cf. Torgerson, 1957). The most frequently used is Guttman's coefficient of reproducibility (R). Depending on the missing data cells, the value of

$$R = 1 - \frac{\text{no. of errors}}{(\text{no. of rows})\ (\text{no. of col.})}$$

R calculated from Table 1 ranges between .964 and .975. Although this is generally considered a very high value, it has been often pointed out that the minimum possible value of R is dependent on the distribution of column sums, so it is desirable to compare obtained values of R with the minimal possible value. In this case, the minimum R would have a value of .678.

Another approach to assessing the fit of the model is to compare the proportion of informants who were found to have perfect response patterns with the proportion of perfect patterns expected if the columns had the same sums, but were completely independent of each other. For any given row response pattern, this is merely the joint product of the probabilities of getting a 1 or 0 on each column. In the present case, the sum of the expected probabilities of the 13

perfect informant response patterns (all 1's to the left of all 0's) is approximately .003. This means that if the features were fully independent of each other, we would expect about three perfect informant response patterns to arise in every 1000 informants observed.[3] Out of the 23 informants in the present study no fewer than 15 had perfect patterns, with a maximum of 17 (depending on how the missing data cells might be assigned as 1's or 0's). The resulting proportions are .65 and .74 respectively as compared with .003 if the items are independent of each other.

These results merely reinforce what can be concluded from even a brief inspection of Table 2, that the informants and linguistic features pattern themselves in a way which is highly systematic and closely fitted by the measurement model of scalogram analysis.

Assigning each informant a Dialect Score (DS) equal to the number of SCTE forms in the perfect response pattern that he exhibits is closest to, we can assume this to be an index of the degree to which his language is standard or nonstandard, and correlate it with indices of other social and psychological variables. For example, we correlated DS with highest grade reached in formal schooling and also with a measure of social-economic class (see Appendix A), abbreviated ISS. Both of these variables are listed for each informant in Table 1. The

---

[3]It should be remembered that there are 2^{12} or 4096 different combinatorially possible response patterns.

correlation between DS and highest grade reached was .80 (Spearman rank-order correlation coefficient) and between DS and ISS it was + .77 ($P < .01$, one-tailed, in both cases). This, of course, merely indicates that the approximation to SCTE spoken by these informants is highly predictable from the amount of education they have had, their present occupation, etc.

Discussion

The results of this study confirm the assumptions made at the beginning of this report. First, we found that different speakers spoke various approximations of SCTE, and thus, that it may not be accurate to characterize a population as consisting of two subgroups--that speaking the standard dialect and that speaking the nonstandard dialect. This is not surprising since most sociological classifications are essentially continuous in nature (for example, SEC). It should be noted that our data do not offer conclusive evidence for or against an underlying linguistic continuum from nonstandard to standard dialects; our sampling of the population was in fact done to emphasize the representation of the ends of the continuum--six of the 23 informants were school teachers and four were functionally illiterate. Also, the question of the existence of an underlying continuum is a theoretical psycholinguistic one which is not necessarily confirmed or refuted by

the observation of a manifest continuum of linguistic behavior. We will come back to this question again below.

Given that a manifest continuum has been observed, we can say that both informants and features can be ordered along it. In particular, it seems that each dialect feature divides the continuum at a particular point; all informants to one side of that point use the nonstandard form of the feature and all those of the other side use the other form. This sort of situation has several convenient implications. If we find an informant displaying a given nonstandard form, we immediately can predict that he will also display the nonstandard forms of all features with cutting points farther toward the standard pole. Thus, if we hear an informant use the form up 'ere or like 'at, we can predict that he will also use adjectives for adverbs, that he will say wadn't for wasn't, that he will use there is for there are and that he will use that for who in certain contexts. On the other hand, if he only uses got very infrequently as a possessive, we can predict that he will not say ain't, that he will not use double negatives, etc.

Intra- versus Interinformant Continua

While the present results do not directly bear on questions of style and style shifts by individual speakers, we believe that the theoretical relationship between dialect and style is a very close one of a psycholinguistic nature and its explication should be considered

as a primary focus for future research. We would like to propose that style shifts by an individual from more formal to less formal, or vice versa, can be characterized as movement by that individual on the dialect continuum from nonstandard to standard. In other words, if the notion of a continuum from nonstandard to standard speech is a reasonable one (which the present results indicate it is), then the suggestion is to make it serve also as the formal-informal style continuum so often talked about by sociolinguists (cf. Labov, 1966). Under this characterization we might think of a person occupying a segment of the nonstandard-standard continuum, with the "lower" end of his segment representing his most informal style and the "upper" end his most formal style. If this is an accurate notion, then changes in style within a speaker should be fully equivalent to holding style constant and going from speaker to speaker. The latter situation, of course, is the only one explored in the present study. Style was considered to be relatively constant since in all cases the interview situation was very similar. This elicited a relatively formal style on the part of all informants, and, while it is possible that momentary style shifts occurred within an interview, our subjective impression is that this was infrequent and of short duration.

If we extrapolate from the present study to one where the speech situation (and thus the style of speech) is experimentally manipulated within informant on the formal-informal dimension, we would

expect that each informant's speech in each condition would be independently locatable on the standard-nonstandard continuum. In addition we could predict, from knowledge of an informant's position on the continuum in a given style, what features will be affected when he switches to a more formal style (or a less formal one). We hope to begin soon to gather experimental evidence bearing on this formulation of style and its relation to dialect.

The Psycholinguistics of Style-Dialect Shifts

The psychological processes involved in style or dialect shifts within an individual have not been explored at all to date; however, several questions can be asked. First of all, some relatively straightforward descriptive questions can be formulated. For example, is the shift from one style to another similar to a shift from one regional dialect to another (for those individuals who are bidialectal)? Do the same sorts of environmental cues trigger these shifts? Can variations in regional dialect be scaled on a continuum in the same way that nonstandard-standard variations have been scaled in this study?

Second, it is possible to speculate about how dialectal variations are best formulated in a performance model of language acquisition, perception, and production. Labov (1968) has presented a first step toward formalizing dialectal or style variations in a grammar.

He proposes that each rule which represents a dialect variation describe--over a set of speakers and sentences--how often that rule will be used. The present results are roughly consistent with his formulation; however, they suggest that for several such rules, which apparently govern quite separate variations (e.g., syntactic vs. phonological), the rule probabilities may operate independently of each other at a given point on the standard-nonstandard continuum, but some mechanism is apparently needed to specify how these probabilities change as one moves along the continuum. It may be that when these formulations are more thoroughly explored, they will have rather direct suggestions for a performance model.

More Immediate Research Objectives

The present results are based on only a part of the data base which has been collected. With respect to informants, another 20 to 30 people were interviewed during the pretesting phase of Stolz & Legum's project. At least some of the resulting recordings are suitable for inclusion with the corpora used here. Also, there are several variables which have not been subjected to scalogram analysis but which might be. These involve the knowledge by informants of various lexical items which may be regional rather than nonstandard. They were specifically elicited by the inverview schedule and so are available for scaling.

REFERENCES

DeCamp, David. Toward a generative analysis of a post-creole speech continuum. Read at the Conference on Pidgenization and Creolization of Language, Mona, Jamaica, 1968.

Guttmann, Louis. A basis for scaling qualitative data, American Sociological Review, 9:139-150, 1944.

Labov, William. Contraction, deletion and inherent variability in the English copula. Paper given at the meetings of the Linguistic Society of America, Chicago, Illinois, 1967.

Labov, William. The Social Stratification of English in New York City (Washington, D.C.: Center for Applied Linguistics, 1966).

Stolz, Walter, & Legum, Stanley. The role of dialect in the school-socialization of lower class children. Final Report on Head Start Evaluation and Research, Section 5. John Pierce-Jones, Director. Child Development Evaluation and Research Center, The University of Texas at Austin, 1967.

Torgerson, Warren. Theory and Methods of Scaling (New York: Wiley, 1957).

PART B

HAVE AND HAVE GOT AS DIALECT VARIANTS IN CENTRAL TEXAS ENGLISH

(THE PROBLEM OF HAVE AND HAVE GOT)

P A R T B

HAVE AND HAVE GOT AS DIALECT VARIANTS IN CENTRAL TEXAS ENGLISH (THE PROBLEM OF HAVE AND HAVE GOT)

Garland Bills

Introduction

This report is an addendum to the preceding fuller study. It is intended as an example of the analysis of a single dialectal variable: have versus have got in possession sentences. The purpose of the analysis is to identify the differences that occur in the dialectal usage patterns, and to attempt to explain how and why these differences occur.

Methods

Informants

The analyses and conclusions presented in this report are based primarily on the interviews conducted with twelve selected informants. The people of the Leander, Texas, region range along a socio-economic scale from working class to upper-middle class. There are few elements of the traditional upper or lower classes. The Index of Social Status (ISS) set up for the Central Texas area by McGuire and White (1955) has been used in determining the class ranking for the informants of this project. In this report that ranking will be subdivided into three groups:

A. ISS 84 - 67 . . . . . working class (Cedarchoppers)

B. ISS 66 - 47 . . . . . (lower-)middle class

C. ISS 46 - 23 . . . . . (upper-)middle class

The twelve informants used in this study are equally divided among these three groups, four informants to each group. In order to minimize the possibility of influences from other dialect regions of the U.S., these twelve informants were selected on the basis of their being life-long residents of the Central Texas area.

Style

Labov (1966) has demonstrated the importance of style in phonological variations and has adapted specific methods for the control of stylistic variations. However, the data collected for the present study were not sufficiently controlled to provide enough syntactic information in a variety of styles. The amount of nonformal speech occurring in the interviews with some informants is negligible and in some cases practically nonexistent. Therefore, it was decided that in the initial analysis all the free speech would be considered "semiformal," for each informant was conversing in familiar surroundings with a highly-educated, upper-middle class stranger. On this basis the major trends of usage could be explored, and later any exceptional occurrences could be reexamined to determine if stylistic phenomena were involved.

Procedures

The initial step in this study was to collect from the speech of the twelve informants all occurrences of possession sentences in which the verbal element was either *have* or *have got*. It was then necessary to determine the syntactically significant environments in which the *have got* form can and cannot occur. It turns out that the environments where *have got* can occur are tightly restricted.

Generally it is used only in the simple present tense (e.g., I have got a book). It is never used with any aspectual or modal modifications of the verb phrase; only the have form is acceptable in such cases. For example (an asterisk marks an unacceptable sentence in the sense intended):

Past Tense:	I had a book.	*I had got a book.
Current Relevance:	I have had a book.	*I have had got a book.
Emphatic do:	I do have a book.	*I do have got a book.
Modal will:	I will have a book.	*I will have got a book.
Modal may:	I may have a book.	*I may have got a book.

etc.

The have got form is usually not allowed in negatives and interrogatives, the common patterns being I don't have a book and Do you have a book?, respectively. Sometimes, however, the have got form is used in these cases by a few of the working class speakers, e.g., I haven't got a book and Have you got a book?, respectively. Since examples such as these last are rare, for the purposes of this study the environment where have got can be employed has been restricted to affirmative declarative sentences in the simple present tense. In addition, the idiomatic constructions such as have a party, have dinner, etc., are also excluded since these constructions never admit the have got form and since they have the peculiarity of allowing the progressive aspect

(we are having lunch), which true possession sentences never allow (*we are having a book).

Results

The possession sentences appearing in the speech of the twelve informants were analyzed to determine the number (N) of sentences in which the have got form might possibly occur. Percentages of the occurrence of have got relative to N were then computed. This information is contained in Table 1, along with the ISS ranking of each informant.

As Table 1 clearly indicates, the speakers of Group A show a preponderance of usage of the have got form, while the informants of Group C tend to use the have form most frequently. However, the speakers of Group B show considerable variation, with some overlapping with both the higher and lower groups. In spite of this variation exhibited by Group B, there clearly exists a definite correlation between socio-economic class and the use of the have/have got variable. The important factor is that this variable shows a bimodal distribution polarized in the two groups, A and C. The form preferred by the upper-middle class (have) we will consider "standard," a feature of Standard Central Texas English (SCTE). The form preferred by the working class (have got) we will call "nonstandard," a characteristic of Nonstandard Central Texas English (NCTE).

TABLE 1

PERCENTAGE OF HAVE GOT AS POSSESSION VERB

Group	Informant	ISS	No. have got	No. have	N	% have got
A	HM	76	15	0	15	100
	EH	80	9	2	11	82
	JT	80	7	3	10	70
	JC	71	7	0	7	100
B	GF	65	5	1	6	83
	DW	58.5	1	3	4	25
	NC	57.5	1	15	16	6
	MVW	56	7	9	16	44
C	EW	43	1	25	26	4
	KF	34	2	16	18	11
	GK	32	1	23	24	4
	AF	32	1	33	34	3

Discussion

Dialect Analysis

The linguistic analysis of these possession sentences must provide for the two dialectal forms. And furthermore, the formal descriptive apparatuses for the two forms must be as similar as possible to account for the fact that most speakers of this region are capable of switching from one form to the other with some degree of facility.

In providing the analysis, I will rely greatly on an article by Bach (1967), the most thorough attempt to account for the use of *have* and *be* in English. However, some modifications will be necessary. Bach treats *have* and *be* as transformationally inserted elements that are attached to the auxiliary (Aux) mode where no lexical verb has been selected. However, in Central Texas English the possessive *have* functions exactly like a lexical verb, but whenever *got* is inserted, this *got* then assumes the function of the lexical verb.[1] Consider the following topics of comparison:

(a) In SCTE the possessive *have* never contracts as do *be* and the auxiliary *have*:

I have a book. *I've a book.

---

[1]Darden (1968) has pointed out similar problems with Bach's analysis, but his alternative approach to a solution is different from that given here.

He is a man.	He's a man.
I have worked.	I've worked.

although in NCTE have does contract where the possessive got fills the verb position:

I have got a book.	I've got a book.

(b) In yes-no interrogatives SCTE requires the pro-verb do with possessive have and with lexical verbs, but not with be or the auxiliary have:

Does he have a book?	*Has he a book?
Does he read fiction?	*Reads he fiction?
*Does he be a man?	Is he a man?
*Does he have worked?	Has he worked?

But in NCTE possessive got acts like a lexical verb and the have is treated as an auxiliary (among those few speakers who allow got in yes-no interrogatives; otherwise, the NCTE form is identical to the SCTE form above):

*Does he have got a book?	Has he got a book?

(c) In negatives the pro-verb do is required in SCTE as stated in (b):

He doesn't have a book.	*He hasn't a book.
He doesn't read fiction.	*He readsn't fiction.
*He doesn't be a man.	He isn't a man.
*He doesn't have worked.	He hasn't worked.

and the NCTE requirements and conditions are also the same as given above in (b):

*He doesn't have got a book.

He hasn't got a book.

The grammatical analysis must specify, therefore, that possessive have is a verb in SCTE and that in NCTE possessive got is a verb and the accompanying have is an Aux. I have attempted to work out various formal solutions to this problem, but none seems to account for several related problems of generality in a satisfactory manner, so I will here present a tentative analysis and simply echo Bach's statement (1967, p. 467):

> Under any analysis so far proposed the special behavior of got with have must be taken care of in an ad hoc fashion. My analysis is no better or worse than others in this respect.

Contrary to Bach's approach, I will assume that V (verb) is an obligatory element of the phrase structure rule for VP (verb phrase), but that the V may later be marked as (-Lexical), thus not being filled by a verb form the lexicon. Such a revision will slightly complicate

Bach's transformational rule for nonverbal sentences and some later embedding and deletion rules where have is affected, but these complications are essential. After these various rules have applied providing for the possessive have, there will then be a "get-insertion rule." This rule is contained in the grammar of all English dialects, accounting for such constructions as the following:

Inception-Recurrence:	"They got to making chain saws."
Permissive:	"He gets to go everywhere."
Adjective Verbalization:	"It gets hot here."
Prepositional Phrase Verbalization:	"I got into trouble."
$Passivization_2$:	"He got thrown off the horse."
Reception:	"He gets his money through the mail."
Acquisition:	"He gets his clothes in Austin."

and others. These types of get are used without tense, mood, or aspect restrictions. However, there is one other construction that parallels the possession construction in restrictions. This is the obligation construction (e.g., I have got to go in variation with I have to go). Unlike possessive got, the use of the obligation got apparently does not correlate strongly with socioeconomic class among the informants studied (although there were too few examples in the interviews to demonstrate this conclusively).

In both SCTE and NCTE, then, there exists a get-insertion rule. The part of this rule that inserts get into obligation constructions is severely limited as to tense, mood, and aspect, and this same part also applies to possession sentences in NCTE, but not in SCTE. Rule I is a simplified version of what this rule might look like:

I.	SD:	X	Aux	(+ V,	- Lex)	Y	Z
		1	2		3	4	5
	SC:	3 --→ get					

There will be, of course, various restrictions as to what the segment Y may be composed of. In addition, where the obligation (and in NCTE the possessive) construction is concerned there will be conditions on the composition of Aux.

There will then be a later rule of "have-movement" of the form:

II.	SD:	X	have	(+ V,	- Lex)	Y
		1	2		3	4
	SC:	3 --→ 2				
		2 --→ ∅				

That is, where get has not been attached to the V node by I, then II attaches an adjacent have to the V node. This rule will account for the verbal function of possessive have. Rule II will automatically operate on all possession sentences in SCTE, but only on the non-get possession sentences in NCTE.

In SCTE rule I will also contain the condition that no negative or interrogative element be included if a possession or obligation construction is dealt with. Thus, SCTE will produce such sentences as:

I don't have a dog.

I don't have to go.

Do you have a dog?

Do you have to go?

This negative/interrogative restriction is also generally heeded by NCTE speakers. But, as noted above, a few speakers ignore this restriction. The interesting point is that only those who allow get in negative/interrogative possession sentences also allow get in negative/interrogative obligation sentences:

I haven't got a dog.

I haven't got to go.

Have you got a dog?

Have you got to go?

This type of analysis, therefore, will readily handle the problems of specifying the two dialectal forms. It also indicates that the same mechanism is involved. Indeed, the only difference between the two dialectal grammars is the fact that a rule contains certain conditions that are adhered to in one dialect and ignored in the other.

Dialects and Style

This type of approach and analysis also has interesting implications for the study of style. Although this project has not delved deeply into this aspect of linguistic behavior, a few comments on the subject of style may be added here.

Style may be defined from the point of view of dialect by saying simply that style is dialect. In other words, style-shift is a change from one dialect to another. Dealing with the single variable discussed in this report we could say that when a speaker uses have he is speaking in a formal (standard) style, and when he uses have got he is speaking in a nonformal (nonstandard) style. Such an approach equates style and dialect and says in effect that nonformal SCTE is NCTE and, vice versa, that formal NCTE is SCTE.

This theoretical approach permits a factor of explanation to be considered in questions of style. That is, it predicts that any time a Central Texas speaker uses have he is speaking formally and when he uses have got he is speaking informally. Formalizing this concept, we might say that the restriction for rule I's application to possession sentences is marked as (+ Formal), i.e., the restriction is heeded only in formal speech.

This viewpoint is readily corroborated in some cases. For example, EW's single example of have got occurred as a statement to his mother when she entered the room (You've got company.). AF's sole use

of have got appeared in a highly excited monologue in which she was explaining the teaching of new math. In both cases the contexts are those that Labov (1966) would consider nonformal speech. However, there is a major problem in dealing with subjects such as MVW who used have got 44% of the time. Deciding when MVW is speaking formally or nonformally is in many cases little more than guess work.

On the other hand, the only alternative approach that can be seen is to accept a formalization of the traditional "free variation" approach to explanation. That is, we could say that the possession sentence restriction of rule I is marked as (+ 56%) for MVW (producing a random 44% have got possession sentences), (+97%) for AF, (+ 30%) for JT, etc. This approach would be able to account for the data in any given text, but it would not provide any real explanatory power. It is probably safe to say that the context of an interview could be so established that MVW would produce 0% have got or that AF would produce more than 50% have got. The approach suggested here offers the possibility of a full explanation in addition to a description, and it would seem to be a fruitful field of future endeavor.

Another problem inherent in the approach that has been advocated is the matter of how to explain the fact that a speaker might consistently give the standard form of one variable and the nonstandard form of another variable. This point was mentioned by Fischer (1958), who suggested that a "formality index" might be established for any

given speaker. This indexing is essentially what has been done in the main body of this report and has offered interesting and consistent results. Apparently, however, formalization of this process would require that each variable be separately indexed as to degree of formality. In the speech of just one person this would add up to a tremendous number of indexing features. The feasibility of index grouping and simplification might be explored as an answer to this problem.

Summary

This addendum report has demonstrated that there exists a definite correlation between the use of _have got_ and socioeconomic class, pointing out that _have_ is a characteristic of SCTE and _have got_ of NCTE. An attempt was made to show how both forms are generated on the basis of a series of rules common to English grammar in general, and to show that minor restrictions on these rules produce the different dialectal forms. In addition, the question of style was explored in relation to a concept of language equating style and dialect, and problems and proposals for future consideration were discussed.

REFERENCES

Bach, Emmon. Have and be in English syntax, Language, 43:462-85, 1967.

Darden, Bill J. Is the English perfect an embedded past? A statement from the devil's advocate. In Papers from the Fourth Regional Meeting Chicago Linguistic Society (Chicago: University of Chicago, 1968).

Fischer, John L. Social influences on the choice of a linguistic variant, Word, 14:47-56, 1958.

Labov, William. The social stratification of English in New York City (Washington, D.C.: Center for Applied Linguistics, 1966).

McGuire, Carson, and George D. White. The estimation of social status. Research Paper in Human Development No. 3 (revised), Department of Educational Psychology, The University of Texas, 1955.

APPENDIX A

APPENDIX A

THE INDEX OF SOCIAL STATUS

The measure of socioeconomic status which was used in this study was originally developed by McGuire and White (1955). It was chosen because it had been developed for use with Texas cities and towns. The measure is called the Index of Social Status (ISS) and is the weighted sum of three rating scales as shown below.

Scale	Range	Weight	Minimum Score	Maximum Score
Occupation	1-7	5	5	35
Source of Income	1-7	4	4	28
Education	1-7	3	3	21
TOTAL			12	84

A low ISS score indicates a high social class and a high ISS score indicates a low social class. Rough equivalents to the usual social class terminology are:

Upper-middle = 23-37

Lower-middle = 38-51

Upper-lower = 52-66

Lower-lower = 67-84

All informants were rated by the interviewer.

The specific rating scales are given below:

APPENDIX A (Continued)

OCCUPATIONS: LEVELS AND KINDS

Rate	Professionals	Proprietors	Businessmen	White Collar	Blue Collar	Service	Farm People
1.	Lawyer, judge, physician, engineer, professor, school supt. et al.	Large business valued at $100,000 or more depending on community	Top executives, President, et al. of corporations, banks, public utilities	CPA: editor of newspaper, magazine; executive secy. of status organization			Gentleman farmer or land owners who do not supervise directly their property
2.	Nurses, teachers, librarians, and others with 4-yr. college degree	Business valued at $50,000 to $100,000	Asst., office, & dept. manager or supervisors; some mfg. agents	Accountant; insurance, real estate, stock salesmen; editorial writers			Land Operators who supervise properties and have an active urban life
3.	Professionals without 4-yr. college degree	Business or equity valued from $10,000 to $50,000	Managers of small branches or buyers and salesmen of known mchdse.	Bank clerks, auto salesmen, postal clerks, RR or Tel. agent or supervisor.	Small contractor who works at or supervises his jobs		Farm owners with "hired help"; operators of leased property who suprvs.
4.		Business or equity valued from $5,000 to $10,000	(Stenographer, bookkeeper; ticket agent, sales people in dept. stores, et al.)		Foreman; master carpenter, electrician, et al., RR engineer	Police capt. tailor, RR cond., watchmkr.	Small landowner; operators of rented property hiring "hands"
5.		Business or equity valued from $2,000 to $5,000	(Dime store clerks, grocery clerks; telephone and beauty oper., et al.)		Apprentice to skilled trades repairmen; med. skilled workers	Policemen; barbers; LVN's, brakemen	Tenants on good farms; foreman; owners of farms who "hire out"
6.		Business or equity valued at less than $2,000		(Semiskilled factory and production workers; assistants to skilled trade; warehousemen, watchmen)		Taxi and trk. drivers; waiter, waitress, gas stn. attnt., aides	Sharecroppers; established farm laborers; subs'ce farmers
7.	"Reputed Lawbreakers"			(Heavy labor; odd-job men; mine or mill hands, unskilled workers)		Domestic hlp., busboy, scrubwomen, janitor help	Migrant workers "squatters and nesters"

APPENDIX A (Continued)

EDUCATION ATTAINMENT*

1. Completed appropriate graduate work for a recognized profession at highest level; graduate of a generally recognized, high status, four-year college.

2. Graduate from a four-year college, university, or professional school with a recognized bachelor's degree, including four-year teacher colleges.

3. Attended college or university for two or more years; junior college graduate; teacher education from a normal school; R.N. from a nursing school.

4. Graduate from high school or completed equivalent secondary education; includes various kinds of "post-high" business education or trade school study.

5. Attended high school, completed grade nine, but did not graduate from high school; for persons born prior to 1900, grade eight completed.

6. Completed grade eight but did not attend beyond grade nine; for persons born prior to 1900, grades four to seven would be equivalent.

7. Left elementary or junior high school before completing grade eight; for persons born prior to 1900, no education or attendance to grade three.

SOURCE OF INCOME**

1. Inherited saving and investments; "old money" reputed to provide basic income.

2. Earned wealth; "new money" has provided "transferable" investment income.

3. Profits, fees, royalties, includes executives who receive a "share of profit."

4. Salary, commissions, regular income aid on monthly or yearly basis.

5. Wages on hourly basis; piece-work; weekly checks as distinguished from monthly.

6. Income from "odd jobs" or private relief; "sharecropping" or seasonal work.

7. Public relief or charity; nonrespectable incomes (reputation).

*Actual education attained probably is not as important as the education a person is reputed to have. The same scale is used to rate aspiration.

**The kind of income appears to be more important than the amount and, in general, the reputed major source of income is symbolic of placement in the community. In the case of a widow, the SI and CC are that of the deceased husband. Investments, insurance, pensions, security benefits, et al. are rated by the SI which made them possible unless considerable wealth ("1" and "2") is reputed. Other components correct for seeming discrepancies.